JOURNAL OF DIFFERENTIAL EQUATIONS 13, 127–149 (1973)

An Attainable Sets Approach to Optimal Control of Functional Differential Equations with Function Space Terminal Conditions

H. T. BANKS*

Center for Dynamical Systems, Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912

AND

MARC Q. JACOBS†

Department of Mathematics, University of Missouri, Columbia, Missouri 65201

Received March 22, 1972

1. INTRODUCTION

In recent papers [2, 9] necessary conditions were derived for optimal control problems involving functional differential systems where the trajectories must satisfy function space terminal conditions such as $x(\theta) = \zeta(\theta)$, $\theta \in [t_1 - h, t_1]$, ζ being some given function. Jacobs and Kao used the Lagrange Multiplier Rule [13] in function space to obtain necessary conditions for optimality for problems with general nonlinear retarded systems while Banks and Kent arrived at equivalent conditions for problems concerned with a wide class of nonlinear systems of neutral type by employing the abstract extremal approach due to Neustadt [15]. In both papers it was shown that for *normal* problems involving linear systems, the necessary conditions are also sufficient in case certain convexity hypotheses are satisfied. The conditions obtained by Banks and Kent were applicable to problems with very general restraints on the controls ($u(t) \in U(t)$) but no results at all concerning normality were discussed in detail. A number of examples utilizing the sufficiency results were presented, demonstrating that the class of normal problems is nonvacuous. Jacobs and Kao were able to establish normality for a class of problems with unrestrained controls ($u(t) \in R^m$), but only under rather severe assumptions on the systems. In this paper we show

* This research was supported in part by the National Aeronautics and Space Administration under Grant No. NGL 40-002-015, and in part by the United States Army, Durham under Grant No. DA-31-124-ARO-D-270.

† This research was supported by the National Science Foundation under Grant No. GP-33882.

that, under relatively weak assumptions on the systems (see H1–H3 in Section 2), one can derive the above-mentioned necessary conditions with normality assured for a wide class of control problems for neutral systems

$$\dot{x}(t) = A_1(t)\,\dot{x}(t-h) + A_2(t)\,x(t) + A_3(t)\,x(t-h) + B(t)\,u(t) \tag{1.1}$$

with unconstrained L_2 controls.

In Section 2 general notation and problem formulation for systems (1.1) are detailed. The idea of the attainable set in function space and its augmented form and saturation are introduced. For the problems considered here, these are natural generalizations of well-known and often used concepts [12]. Under the hypotheses H1–H3 it is shown in Section 3 that the attainable set is a closed linear variety in $W_2^{(1)}$. This enables one to conclude that the saturation of the augmented attainable set is closed convex with nonempty interior relative to an appropriately chosen Hilbert space. In this Hilbert space one can establish the existence of the needed supporting hyperplanes through the boundary points of the saturation of the augmented attainable set. Then necessary conditions for integral quadratic cost problems can be derived using standard geometric ideas. This is done in Theorem 4.1, where normality is guaranteed. In fact, it is shown in Theorem 4.2 that one can actually obtain necessary conditions with normality for problems with general Lagrange costs and systems (1.1) by applying the Lagrange Multiplier Rule cited above.

The hypotheses H1–H3 arise out of an attempt to show that the saturation of the augmented attainable set, $\mathrm{sat}(\hat{\mathscr{A}}_0)$, has nonempty interior relative to a certain subspace of $R \times W_2^{(1)}$. As is pointed out in Remark 3.1 below, a necessary condition for $\mathrm{sat}(\hat{\mathscr{A}}_0)$ to have nonempty interior in $R \times W_2^{(1)}$ is that the rank of $B(t)$ be n a.e. on $[t_1 - h, t_1]$. Strengthened forms of this necessary condition (e.g., B continuous and $B(t)$ has rank n at *each* t in $[t_1 - h, t_1]$) can easily be shown to imply $\mathrm{sat}(\hat{\mathscr{A}}_0)$ has nonempty interior relative to $R \times W_2^{(1)}$.

In this paper we shall omit all discussions concerning motivation for the problems studied here. Comments on that aspect of these problems, along with a number of solved examples, can be found in [1, 2, 9]. Moreover, in [1] one will find a general survey of results and techniques involved in questions of controllability, existence of optimal controls, and necessary and sufficient conditions for optimality for control problems with function space boundary conditions and functional differential equation systems.

2. Notation and Formulation of the Optimal Control Problems

The symbol $\mathscr{L}_{pq}$ denotes the collection of all $p \times q$ real matrices with a suitable matrix norm. The functions $t \mapsto A_i(t)$, $i = 1, 2, 3$ and $t \mapsto B(t)$

in (1.1) which map R (the set of real numbers) into $\mathscr{L}_{nn}$ and $\mathscr{L}_{nm}$, respectively, are assumed to be continuous. In addition A_1 is assumed to be continuously differentiable on R. The symbol $L_2([a, b], R^p)$ denotes the usual Hilbert space of "square integrable" functions from $[a, b]$ into R^p [5]. In all cases where the notion of measure intervenes Lebesgue measure is understood unless stated expressly to the contrary. We use $W_2^{(1)}([a, b], R^p)$ to denote the Sobolev space consisting of all absolutely continuous functions $x: [a, b] \to R^p$ with the property that the function $t \mapsto \dot{x}(t) = (dx/dt)(t)$ belongs to $L_2([a, b], R^p)$. The space $W_2^{(1)}([a, b], R^p)$ is a Hilbert space with inner product defined by

$$\langle x, y\rangle \equiv \langle x(a), y(a)\rangle + \int_a^b \langle \dot{x}(t), \dot{y}(t)\rangle\, dt,$$

for $x, y \in W_2^{(1)}([a, b], R^p)$ (the angle brackets on the right stand for the usual inner product of vectors in R^p). All vectors $x \in R^p$ will be viewed as column vectors ($p \times 1$ matrices), and we use A^* to denote the transpose of a matrix $A \in \mathscr{L}_{pq}$.

Suppose $h > 0$ is the time lag in system (1.1) and $x: [t_0 - h, t_1] \to R^n$ Then for $t \in [t_0, t_1]$, x_t denotes the function on $[-h, 0]$ defined by

$$x_t(\theta) = x(t + \theta), \qquad -h \leqslant \theta \leqslant 0.$$

If $\varphi \in W_2^{(1)}([-h, 0], R^n)$ and $u \in L_2([t_0, t_1], R^m)$ are given, then $x(\cdot, \varphi, u)$ is the response function, $t \mapsto x(t, \varphi, u)$, $t_0 \leqslant t \leqslant t_1$, to (1.1) with control u and initial condition

$$x_{t_0} = \varphi. \tag{2.1}$$

Let $L: R \times R^n \times R^m \to R$ be a given continuous function, and define

$$J(x, u) = \int_{t_0}^{t_1} L(t, x, u)\, dt. \tag{2.2}$$

Let ζ be another function in $W_2^{(1)}([-h, 0], R^n)$. The type of optimization problems which we shall examine is that of minimizing (globally) $J(x, u)$ on the class of all controllers $u \in L_2([t_0, t_1], R^m)$ such that (2.1), (1.1),

$$x_{t_1} = \zeta, \tag{2.3}$$

and $|J(x, u)| < \infty$ are satisfied. The time interval $[t_0, t_1]$ with $t_1 > t_0 + h$ will be fixed in our discussion.

The *attainable set at time t_1 starting from initial function* φ is defined by

$$\mathscr{A}_\varphi = \{x_{t_1}(\cdot, \varphi, u) \mid u \in L_2([t_0, t_1], R^m)\}. \tag{2.4}$$

The corresponding *augmented attainable set* is given by

$$\hat{\mathscr{A}}_\varphi = \{(x^0, x_{t_1}(\cdot, \varphi, u)) \mid u \in L_2([t_0, t_1], R^m),\\ x^0 = J(x(\cdot, \varphi, u), u), \qquad |x^0| < \infty\}. \tag{2.5}$$

The saturation of $\hat{\mathscr{A}}_\varphi$ is defined by

$$\text{sat}(\hat{\mathscr{A}}_\varphi) = \{(y^0, x_{t_1}) \mid y^0 \geqslant x^0, (x^0, x_{t_1}) \in \hat{\mathscr{A}}_\varphi\}. \tag{2.6}$$

Let $X(t, s)$ be the unique $n \times n$ matrix solution to the following integral equation

$$X(t, s) = I + X(t, s+h)A_1(s+h)\\ + \int_s^t X(t, \alpha)A_2(\alpha)\, d\alpha + \int_{s+h}^t X(t, \alpha)A_3(\alpha)\, d\alpha \tag{2.7}$$

for $t_0 \leqslant s < t \leqslant t_1$ subject to conditions

$$X(t, t) = I, \qquad X(t, s) = 0 \qquad \text{for} \quad s > t, \tag{2.8}$$

where I is the $n \times n$ identity matrix. The matrix function X is of bounded variation on $[t_0, t_1]$ in both variables separately. For discussion of this and other properties of X see [2, 3, 7]. The variation of constants formula for (1.1) gives

$$x(t, \varphi, u) = x(t, \varphi, 0) + \int_{t_0}^t X(t, s)B(s)u(s)\, ds, \qquad t_0 \leqslant t \leqslant t_1. \tag{2.9}$$

We now state some hypotheses that will be referred to constantly in the sequel.

(H1) $G(t_0, t_1 - h) \equiv \int_{t_0}^{t_1-h} X(t_1 - h, s)B(s)B^*(s)X^*(t_1 - h, s)\, ds$ has rank n.

(H2) There exist bounded measurable matrix functions Γ_1, Γ_3: $[t_1 - h, t_1] \to \mathscr{L}_{mn}$ such that

$$\left.\begin{aligned} A_1(t) &= B(t)\Gamma_1(t) \\ A_3(t) &= B(t)\Gamma_3(t) \end{aligned}\right\}, \quad t_1 - h \leqslant t \leqslant t_1.$$

(H3) $B^\dagger(t)$, the generalized (or pseudo-) inverse of $B(t)$ (see [14, or 16]), is bounded on $[t_1 - h, t_1]$.

Remark 2.1. It is noted that if $B^*(t)\, B(t)$ is invertible for each $t \in [t_1 - h, t_1]$, then $B^\dagger(t) = [B^*(t)\, B(t)]^{-1}B^*(t)$ for $t \in [t_1 - h, t_1]$. On the other hand, if

$B(t)\,B^*(t)$ is invertible for each $t \in [t_1 - h, t_1]$, then $B^\dagger(t) = B^*(t)[B(t)\,B^*(t)]^{-1}$. Thus in either of these cases (H3) is a consequence of the continuity of $t \mapsto B(t)$. Of the two situations, the former often proves to be the most useful (cf. the example in Section 3).

Remark 2.2. The above two situations are special cases of instances where $B(t)$ has constant rank for $t \in [t_1 - h, t_1]$. In all cases where $B(t)$ has this property, hypothesis (H3) is satisfied. This follows from the continuity of $t \mapsto B(t)$ plus the fact that the mapping $C \mapsto C^\dagger$ of $\mathscr{L}_{nm}$ into $\mathscr{L}_{mn}$ is continuous on regions of $\mathscr{L}_{nm}$ where the rank is constant [16].

3. A Closure Lemma

In this section we establish the following fundamental closure lemma.

Lemma 3.1. *If* (H1)–(H3) *are satisfied, then* $\mathscr{A}_0$ *is a closed linear manifold in* $W_2^{(1)}([-h, 0], R^n)$.

Proof. Consistent with the notation in (2.4) $\mathscr{A}_0$ represents the set in (2.4) with $\varphi \equiv 0$. It is easy to show that $\mathscr{A}_0 \subset W_2^{(1)}([-h, 0], R^n)$. The variation of constants formula (2.9) clearly implies $\mathscr{A}_0$ is a linear manifold. Thus we need only prove that $\mathscr{A}_0$ is closed in $W_2^{(1)}([-h, 0], R^n)$. Let p be a positive integer. Define

$$\mathscr{A}_0{}^p = \{x_{t_1}(\cdot, 0, u) \mid u \in L_2([t_0\,, t_1], R^m), \| u \|_2 \leqslant p\},$$

where $\| u \|_2$ denotes the norm of u in $L_2([t_0\,, t_1], R^m)$. A very simple argument using the weak compactness (in L_2) of the set $\{u \in L_2([t_0\,, t_1], R^m) \mid \| u \|_2 \leqslant p\}$ and the variation of constants formula will show that

$$\mathscr{A}_0{}^p \text{ is a closed subset of } W_2^{(1)}([-h, 0], R^n), \qquad p = 1, 2, 3,\ldots . \tag{3.1}$$

Now suppose that $y^\nu = x_{t_1}(\cdot, 0, u^\nu) \in \mathscr{A}_0$, $\nu = 1, 2, 3,\ldots$ and $y^\nu \to y$ in $W_2^{(1)}([-h, 0], R^n)$. We must show that there is a $u \in L_2([t_0\,, t_1], R^m)$ such that $y = x_{t_1}(\cdot, 0, u)$. Let x^ν be an abbreviation for the function $t \mapsto x(t, 0, u^\nu)$, $t_0 - h \leqslant t \leqslant t_1$. Define a control $\tilde{u}^\nu$ on the interval $[t_0\,, t_1 - h]$ by the equation

$$\tilde{u}^\nu(t) = B^*(t)\,X^*(t_1 - h, t)\,\xi^\nu, \qquad t_0 \leqslant t \leqslant t_1 - h, \tag{3.2}$$

where $\xi^\nu \in R^n$ is the solution to

$$G(t_0\,, t_1 - h)\,\xi^\nu = x^\nu(t_1 - h).$$

This equation has a unique solution by hypothesis (H1). Moreover, since the sequence $x^\nu(t_1 - h)$ converges, the same can be said for the sequence ξ^ν. It follows at once from (2.9) that

$$x(t_1 - h, 0, \tilde{u}^\nu) = x^\nu(t_1 - h). \tag{3.3}$$

Also since the sequence ξ^ν is bounded, the sequence $\tilde{u}^\nu$ is bounded in $L_2([t_0, t_1 - h], R^m)$. Observe that

$$\dot{x}^\nu(t) - A_2(t)\, x^\nu(t) = B(t)[\Gamma_1(t)\, \dot{x}^\nu(t - h) + \Gamma_3(t)\, x^\nu(t - h) + u^\nu(t)] \tag{3.4}$$

a.e. on $[t_0, t_1]$ by hypothesis (H2) applied to the differential Eq. (1.1). Define

$$p^\nu(t) = \Gamma_1(t)\, \dot{x}^\nu(t - h) + \Gamma_3(t)\, x^\nu(t - h) + u^\nu(t) \qquad \text{a.e. on } [t_1 - h, t_1].$$

From (3.4) we have that

$$B(t)\, p = \dot{x}^\nu(t) - A_2(t)\, x^\nu(t) \tag{3.5}$$

has a solution $p^\nu(t)$ a.e. on $[t_1 - h, t_1]$ so that $\dot{x}^\nu(t) - A_2(t)\, x^\nu(t)$ is in the range space of the linear transformation $B(t)$ a.e. on $[t_1 - h, t_1]$. Consequently,

$$\tilde{p}^\nu(t) \equiv B^\dagger(t)[\dot{x}^\nu(t) - A_2(t)\, x^\nu(t)] \tag{3.6}$$

$t_1 - h \leqslant t \leqslant t_1$ is also a solution to (3.5) a.e. on $[t_1 - h, t_1]$. By hypothesis $y^\nu = x_{t_1}(\cdot, 0, u^\nu) \to y$ in $W_2^{(1)}([-h, 0], R^n)$ so that (3.6) and hypothesis (H3) imply that $\tilde{p}^\nu$ is a bounded sequence in $L_2([t_1 - h, t_1], R^m)$. Let $\tilde{x}^\nu = x(\cdot, 0, \tilde{u}^\nu)$ on $[t_0, t_1 - h]$ where $\tilde{u}^\nu$ is given in (3.2) on the interval $[t_0, t_1 - h]$. Extend the control $\tilde{u}^\nu$ to the entire interval $[t_0, t_1]$ by defining

$$\tilde{u}^\nu(t) \equiv \tilde{p}^\nu(t) - \Gamma_1(t)\, \dot{\tilde{x}}^\nu(t - h) - \Gamma_3(t)\, \tilde{x}^\nu(t - h) \tag{3.7}$$

for $t_1 - h < t \leqslant t_1$. Since $\tilde{u}^\nu$ is a bounded sequence in $L_2([t_0, t_1 - h], R^m)$, hypothesis (H2), the variation of constants formula (2.9), and the fact that $\tilde{p}^\nu$ is a bounded sequence in L_2-norm imply that the sequence of functions $\tilde{u}^\nu$ defined on $[t_0, t_1]$ by (3.2) and (3.7) is a bounded sequence in $L_2([t_0, t_1], R^m)$. Moreover, $\tilde{x}^\nu$ is also extended to all of $[t_0, t_1]$ by defining $\tilde{x}^\nu = x(\cdot, 0, \tilde{u}^\nu)$, and the construction above gives $\tilde{x}^\nu_{t_1} = y^\nu$. The situation now reverts to statement (3.1) to give immediately the existence of a $u \in L_2([t_0, t_1], R^m)$ such that $x_{t_1}(\cdot, 0, u) = y$. This proves the lemma.

LEMMA 3.2. *Let Q: $[t_0, t_1] \to \mathscr{L}_{nn}$, and N: $[t_0, t_1] \to \mathscr{L}_{mm}$ be continuous. Let $Q(t)$ be symmetric and positive semidefinite, and let $N(t)$ be symmetric and positive definite for each $t \in [t_0, t_1]$. Let $L(t, x, u) \equiv \langle x, Q(t)x \rangle + \langle u, N(t)u \rangle$, $t \in [t_0, t_1]$, $x \in R^n$, $u \in R^m$. Then* sat$(\hat{\mathscr{A}}_0)$ *is a closed convex subset of*

$R \times W_2^{(1)}([-h, 0], R^n)$. *Moreover,* $\text{sat}(\hat{\mathscr{A}}_0)$ *has a nonempty interior relative to* $R \times \mathscr{A}_0$ *if* (H1)–(H3) *are satisfied.*

Proof. The convexity of $\text{sat}(\hat{\mathscr{A}}_0)$ is an easy consequence of the convexity of the function L in x, u, and the variation of constants formula (2.9). A weak compactness argument similar to that given in [12, p. 232] will establish that $\text{sat}(\hat{\mathscr{A}}_0)$ is closed in $W_2^{(1)}([-h, 0], R^n)$. The details are omitted. Now let hypotheses (H1)–(H3) be satisfied. Let B_ϵ, $\epsilon > 0$, be the open ball with radius ϵ and center 0 in $W_2^{(1)}([-h, 0], R^n)$. Let $B_\epsilon(\mathscr{A}_0) \equiv B_\epsilon \cap \mathscr{A}_0$. It is not difficult to establish that there is a constant $M = M(\epsilon) > 0$ such that if $y \in B_\epsilon(\mathscr{A}_0)$ there is a $\tilde{u} \in L_2([t_0, t_1], R^m)$ whose L_2-norm does not exceed M such that $y = x_{t_1}(\cdot, 0, \tilde{u})$. The demonstration that such an M exists and the construction of $\tilde{u}$ is entirely similar to the construction of the sequence $\tilde{u}^\nu$ in (3.2) and (3.7) and will not be repeated here. A simple calculation now shows that there is an $H = H_\epsilon$ such that

$$0 \leqslant J(x(\cdot, 0, \tilde{u}), \tilde{u}) \leqslant H$$

for $y = x_{t_1}(\cdot, 0, \tilde{u}) \in B_\epsilon(\mathscr{A}_0)$ and the $\tilde{u}$ is as above with L_2-norm not exceeding M. The constant H depends only on $\epsilon > 0$ and $M(\epsilon)$. Hence since the function $x_{t_1} \equiv 0$ is in $\mathscr{A}_0$ we have that

$$(H, \infty) \times B_\epsilon(\mathscr{A}_0) \subset \text{sat}(\hat{\mathscr{A}}_0)$$

and the lemma is proved.

Remark 3.1. We shall give a geometric proof of the maximum principle for the optimization problem formulated in Section 2 with quadratic cost functional. To do this one needs to know when the closed convex set $\text{sat}(\hat{\mathscr{A}}_0)$ has the property that through every boundary point there is a support hyperplane for $\text{sat}(\hat{\mathscr{A}}_0)$. Klee [11] has given an example which shows that this cannot be expected in general. For example, $C = \{\xi = \{\xi^k\} \mid \xi \in l_2, \xi^k \geqslant 0 \,\forall k\}$ is a closed convex set in l_2 in which each point is a boundary point. But C is supported only at points having at least one zero coordinate. When the hypotheses (H1)–(H3) are imposed, then $\text{sat}(\hat{\mathscr{A}}_0)$ has nonempty interior relative to $R \times \mathscr{A}_0$. Since $R \times \mathscr{A}_0$ is a closed subspace of $R \times W_2^{(1)}([-h, 0], R^n)$ by Lemma 3.1, it is also a Hilbert space with the inner product inherited from $R \times W_2^{(1)}([-h, 0], R^n)$. Thus $\text{sat}(\hat{\mathscr{A}}_0)$ has a support hyperplane through each relative boundary point [13, p. 133]. Since $R \times \mathscr{A}_0$ is self-dual the linear functional representing a given support hyperplane for $\text{sat}(\hat{\mathscr{A}}_0)$ is merely an element of $R \times \mathscr{A}_0$.

Remark 3.2. In general, the interior of $\text{sat}(\hat{\mathscr{A}}_0)$ in $R \times W_2^{(1)}([-h, 0], R^n)$ is empty. If the interior of $\text{sat}(\hat{\mathscr{A}}_0)$ is not empty, then it is easily shown that $\mathscr{A}_0$

has nonempty interior in $W_2^{(1)}([-h, 0], R^n)$. Since $\mathscr{A}_0$ is a linear manifold, we have that $\mathscr{A}_0 = W_2^{(1)}([-h, 0], R^n)$. This has very restrictive implications. That is, given any $y \in W_2^{(1)}([-h, 0], R^n)$, there is a $u \in L_2([t_0, t_1], R^m)$ such that (1.1) is satisfied with boundary conditions

$$x_{t_0} = 0, \qquad x_{t_1} = y.$$

Thus in the case of retarded systems ($A_1 \equiv 0$) we have

$$\dot{y}(t - t_1) - A_2(t)\, y(t - t_1) = A_3(t)\, x(t - h) + B(t)\, u(t) \tag{3.8}$$

a.e. on $[t_1 - h, t_1]$. Since $y \in W_2^{(1)}([-h, 0], R^n)$ is arbitrary, the left hand side of (3.8) can be any function in $L_2([t_1 - h, t_1], R^n)$. The function u and hence the function $t \mapsto x(t - h)$ (which depends on u on the interval $[t_0, t_1 - h]$) are dependent on the choice of y. By a simple argument it can be shown in a manner entirely similar to the proof of Lemma 3.1 in [9] that the rank of $B(t)$ must be n a.e. on $[t_1 - h, t_1]$. In many practical problems this would be a clearly undesirable assumption.

Remark 3.3. Algebraic criteria to insure that hypothesis (H1) is satisfied are known in the case of retarded equations (e.g., see [10, 18–20]). For neutral equations the following simple observation generates useable criteria which imply (H1). Assume that system (1.1) is autonomous (A_i, $i = 1, 2, 3$, and B do not depend on t). Suppose the ordinary control system

$$\dot{x} = A_2 x + Bu \tag{3.9}$$

is controllable in the usual sense [12]. This is equivalent to assuming

$$\operatorname{Rank}[B, A_2 B, \ldots, A_2^{n-1} B] = n. \tag{3.10}$$

Then condition (H1) is satisfied for the autonomous system

$$\dot{x}(t) = A_1 \dot{x}(t - h) + A_2 x(t) + A_3 x(t - h) + Bu(t). \tag{3.11}$$

In order to see this observe that if $0 < \delta < h$, then

$$\begin{aligned} &G(t_0, t_1 - h) \\ &\quad = G(t_1 - h - \delta, t_1 - h) + \int_{t_0}^{t_1 - h - \delta} X(t_1 - h, s)\, BB^* X^*(t_1 - h, s)\, ds. \end{aligned}$$

Since both terms on the right hand side of this equation are symmetric and positive semidefinite, it follows that $G(t_1 - h - \delta, t_1 - h)$ being positive definite will imply that $G(t_0, t_1 - h)$ is positive definite thereby showing

$G(t_0, t_1 - h)$ has rank n. Thus the question is reduced to verifying $G(t_1 - h - \delta, t_1 - h)$ has rank n. Appealing to (2.7) we find that

$$G(t_1 - h - \delta, t_1 - h) = \int_{t_1-h-\delta}^{t_1-h} X(t_1 - h, s)\, BB^*X(t_1 - h, s)\, ds$$

$$= \int_{t_1-h-\delta}^{t_1-h} e^{A_2(t_1-h-s)}\, BB^* e^{A_2^*(t_1-h-s)}\, ds$$

which has rank n by (3.10) (see [12]).

EXAMPLE 3.1. Let $a_0, a_1, \ldots, a_{n-1}, b_0, b_1, \ldots, b_n$ be given real numbers. Define

$$B \equiv \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \qquad A_1 \equiv \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & b_n \end{pmatrix},$$

$$A_2 \equiv \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ a_0 & a_1 & a_2 & \cdots & a_{n-1} \end{pmatrix}, \qquad A_3 \equiv \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \\ b_0 & b_1 & \cdots & b_{n-1} \end{pmatrix}.$$

Then hypothesis (H2) is validated by choosing

$$\Gamma_1 \equiv (0, 0, \ldots, b_n),$$

$$\Gamma_3 \equiv (b_0, b_1, \ldots, b_{n-1}).$$

Also (3.10) is true for this example so that Remark 3.3 shows that (H1) is satisfied. Moreover $B^*B = 1$ so that Remark 2.1 gives that hypotheses (H3) is fulfilled. Thus we see that all of our results apply to the general n-th order scalar autonomous neutral equations of the form

$$x^{(n)}(t) = \sum_{i=0}^{n} b_i x^{(i)}(t - h) + \sum_{i=0}^{n-1} a_i x^{(i)}(t) + u(t). \tag{3.12}$$

One can extend Halanay's arguments in [6] to show (see [8]) that if ξ, η: $[-h, 0] \to R$ are such that

$$\varphi = (\xi, \dot{\xi}, \ldots, \xi^{(n-1)}) \in W_2^{(1)}([-h, 0], R^n),$$

$$\zeta = (\eta, \dot{\eta}, \ldots, \eta^{(n-1)}) \in W_2^{(1)}([-h, 0], R^n),$$

then there is a control u such that (1.1) (with A_i, $i = 1, 2, 3$, and B as listed above), (2.1), and (2.3) are all satisfied. This can be regarded as an extension of the remarks of Kirillova and Curakova [10] and Halanay [6] pointing out that scalar retarded equations ((3.12) with $b_n = 0$) are always controllable to the zero function.

Remark 3.4. It is clear from the proof of Lemma 3.1 that if $\zeta \in W_2^{(1)}([-h, 0], R^n)$ and $\dot{\zeta}(t - t_1) - A_2(t)\, \zeta(t - t_1)$ is in the image of $B(t)$ for almost every t in $[t_1 - h, t_1]$, and if hypotheses (H1)–(H3) are satisfied, then there is a $u \in L_2([t_0, t_1], R^m)$ such that (1.1), (2.1) and (2.3) are satisfied. It is precisely this situation that obtains in the controllability remarks in Example 3.1.

4. Necessary Conditions for Optimal Control

In this section necessary conditions for the optimization problem posed in Section 2 are obtained.

THEOREM 4.1. *Suppose* $L(t, x, u) = \langle x, Q(t)x\rangle + \langle u, N(t)u\rangle$, $t_0 \leqslant t \leqslant t_1$, $x \in R^n$, $u \in R^m$, *where Q and N are as in Lemma* 3.2. *Let (P) denote the problem of minimizing (globally) the functional J in* (2.2) *subject to* (1.1), (2.1), (2.3), *and* $u \in L_2([t_0, t_1], R^m)$. *Then the following two propositions are valid*:

(A) *Problem* (P) *has at most one solution. Moreover, if there is at least one* $u \in L_2([t_0, t_1], R^m)$ *such that* (1.1), (2.1), *and* (2.3) *are satisfied, then problem* (P) *has at least one solution.*

(B) *Let the hypotheses* (H1)–(H3) *be satisfied. Suppose* $(\bar{x}, \bar{u})$ *is a solution to Problem* (P). *Then there is a square integrable function* $\mu\colon [t_0, t_1] \to R^n$ *which is constant outside of* $[t_1 - h, t_1]$ *such that*

(i) $\partial L/\partial u(t, \bar{x}(t), \bar{u}(t))^* - \psi(t)^* B(t) = 0$ a.e. *on* $[t_0, t_1]$, *where* ψ *is a solution to the adjoint relations,*

$$\text{(ii)}\quad \psi(t)^* = \mu(t)^* + \psi(t+h)^* A_1(t+h) + \int_{t+h}^{t_1} \psi(s)^* A_3(s)\, ds$$

$$+ \int_t^{t_1} [-\frac{\partial L}{\partial x}(s, \bar{x}(s), \bar{u}(s)) + \psi(s)^* A_2(s)]\, ds,\ t_0 \leqslant t \leqslant t_1,$$

$$\psi(t) = 0 \qquad \text{if} \quad t > t_1 .$$

Proof of Theorem 4.1. The questions of existence and uniqueness of the optimal control were discussed in [2, 9]. Thus we consider only the proof of the necessary conditions. Suppose $\bar{u}$, $\bar{x} = x(\cdot, \varphi, \bar{u})$ is a solution to problem

(P). We will consider only the special case where $\varphi \equiv 0$. Since $x(t, \varphi, u) = x(t, \varphi, 0) + x(t, 0, u)$, the general result is obtained by appropriately translating the special case we consider. Thus $\bar{x} = x(\cdot, 0, \bar{u})$. It follows at once that $(\bar{x}^0, \bar{x}_{t_1})$ must belong to the boundary of $\operatorname{sat}(\hat{\mathscr{A}}_0)$ relative to $R \times \mathscr{A}_0$, where

$$\bar{x}^0 \equiv J(\bar{x}, \bar{u}).$$

In view of Lemma 3.2 and the supporting hyperplane theorem [13, p. 133] there is a nonzero element $(\alpha^0, \lambda) \in R \times \mathscr{A}_0$ such that

$$\alpha^0(\bar{x}^0 - x^0) + \langle \bar{x}_{t_1} - x_{t_1}, \lambda \rangle \geqslant 0 \tag{4.1}$$

for each $(x^0, x_{t_1}) \in \hat{\mathscr{A}}_0$. Suppose $\alpha^0 = 0$. Then

$$\langle \bar{x}_{t_1} - x_{t_1}, \lambda \rangle \geqslant 0$$

for every $x_{t_1} \in \mathscr{A}_0$. Since $\mathscr{A}_0$ is a linear space and $\lambda \in \mathscr{A}_0$, then $-\lambda \in \mathscr{A}_0$. So choose x_{t_1} and $\tilde{x}_{t_1}$ such that

$$\bar{x}_{t_1} - x_{t_1} = \lambda, \qquad \bar{x}_{t_1} - \tilde{x}_{t_1} = -\lambda.$$

Then both $\langle \lambda, \lambda \rangle \geqslant 0$ and $-\langle \lambda, \lambda \rangle \geqslant 0$. Hence $\lambda \equiv 0$ and $(\alpha^0, \lambda) = (0, 0)$, a contradiction. Since the scalar $\alpha^0 \neq 0$ we may assume $\alpha^0 = -1$ and (4.1) becomes

$$-(\bar{x}^0 - x^0) + \langle \bar{x}_{t_1} - x_{t_1}, \lambda \rangle \geqslant 0 \tag{4.2}$$

for each $(x^0, x_{t_1}) \in \hat{\mathscr{A}}_0$, where $x_{t_1} = x_{t_1}(\cdot, 0, u)$ for some admissible u. In the rest of this section the zero initial function is understood, unless explicitly stated otherwise. Hence we shall abbreviate $x(t, 0, u)$ with $x(t, u)$. The variation of constants formula (2.9) gives

$$x(t, u) = \int_{t_0}^{t} X(t, s)B(s)u(s)\, ds, \qquad t_0 \leqslant t \leqslant t_1. \tag{4.3}$$

The mapping $u \mapsto x(\cdot, u)$ is a bounded linear operator from $L_2([t_0, t_1], R^m)$ into $W_2^{(1)}([t_0, t_1], R^n)$. Now $u \mapsto x^0 = J(x(\cdot, u), u) \equiv J(u)$ is a Frechet differentiable mapping of $L_2([t_0, t_1], R^m)$ into R. We will use the notation

$$\bar{L}_x(t) = (\partial L/\partial x)(t, \bar{x}(t), \bar{u}(t)),$$

$$\bar{L}_u(t) = (\partial L/\partial u)(t, \bar{x}(t), \bar{u}(t)),$$

noting that in fact

$$\bar{L}_x(t) = 2Q(t)\,\bar{x}(t),$$
$$\bar{L}_u(t) = 2N(t)\,\bar{u}(t).$$

We keep the calculations in terms of the function L to facilitate the proof of the next theorem. Applying (4.3) and Fubini's theorem, it is determined that

$$\begin{aligned} x^0 - \bar{x}^0 &= J(u) - J(\bar{u}) = J'(\bar{u})(v) + o(\| v \|_2) \\ &= \int_{t_0}^{t_1} \{\langle \bar{L}_x(t), x(t, v)\rangle + \langle \bar{L}_u(t), v(t)\rangle\}\, dt + o(\| v \|_2) \\ &= \int_{t_0}^{t_1} \Big[\int_{t_0}^{t_1} \bar{L}_x(t)^* X(t, s)\, dt\Big] B(s) v(s)\, ds \\ &\quad + \int_{t_0}^{t_1} \bar{L}_u(s)^* v(s)\, ds + o(\| v \|_2), \end{aligned} \tag{4.4}$$

where $v \equiv u - \bar{u}$. The second term on the left hand side of (4.2) must also be calculated. From the definition of the inner product in $W_2^{(1)}([-h, 0], R^n)$, we get

$$\begin{aligned} \langle x_{t_1} - \bar{x}_{t_1}, \lambda\rangle &= \langle x(t_1 - h, v), \lambda(-h)\rangle \\ &\quad + \int_{-h}^{0} \langle \frac{dx}{d\theta}(t_1 + \theta, v), \frac{d\lambda}{d\theta}(\theta)\rangle\, d\theta, \end{aligned} \tag{4.5}$$

where again $v \equiv u - \bar{u}$. This can be written

$$\begin{aligned} \langle x_{t_1} - \bar{x}_{t_1}, \lambda\rangle &= \langle x(t_1 - h, v), \lambda(-h)\rangle \\ &\quad + \int_{t_1 - h}^{t_1} \langle \dot{x}(t, v), \dot{\lambda}(t - t_1)\rangle\, dt. \end{aligned} \tag{4.6}$$

The function $x(\cdot, v)$ satisfies the differential equation (1.1) so (4.6) implies

$$\begin{aligned} \langle x_{t_1} - \bar{x}_{t_1}, \lambda\rangle &= \langle x(t_1 - h, v), \lambda(-h)\rangle \\ &\quad + \int_{t_1 - h}^{t_1} \langle \dot{\lambda}(t - t_1), A_1(t)\dot{x}(t - h, v) + A_2(t) x(t, v) \\ &\quad + A_3(t) x(t - h, v) + B(t) v(t)\rangle\, dt. \end{aligned} \tag{4.7}$$

Since λ and v are equal almost everywhere to Borel measurable functions [17, p. 278], we can assume λ and v are Borel measurable. Indeed, for λ we shall abide by the following convention: When we write $\dot\lambda$ we mean that

$$\dot\lambda(t) = \begin{cases} g(t), & t \in [-h, 0] \setminus E_0\,, \\ 0, & t \in E_0\,, \end{cases}$$

where $g(t) \equiv \limsup_{n\to\infty} n[\lambda(t + 1/n) - \lambda(t)]$, $t \in [-h, 0]$, and $E_0 \equiv g^{-1}(\{\infty\})$. Since λ is absolutely continuous on $[-h, 0]$, the set E_0 has measure zero, and $(d\lambda/dt)(t) = g(t)$ at every point where λ is differentiable. The function g is Borel measurable which implies E_0 is a Borel measurable set. Hence $\dot\lambda$ interpreted as above is Borel measurable.

By making an appropriate change of variables in the integral in (4.7) and using (4.3) the integral term in (4.7) can be written as

$$\begin{aligned} &\int_{t_1-2h}^{t_1-h} \left\langle \lambda(t + h - t_1),\, A_1(t + h)d_t \left[\int_{t_0}^{t} X(t, s)B(s)v(s)\, ds\right]\right\rangle \\ &\quad + \int_{t_1-2h}^{t_1-h} \left\langle \dot\lambda(t + h - t_1),\, A_3(t + h) \int_{t_0}^{t} X(t, s)B(s)v(s)\, ds\right\rangle dt \\ &\quad + \int_{t_1-h}^{t_1} \left\langle \lambda(t - t_1),\, \left[A_2(t) \int_{t_0}^{t} X(t, s)B(s)v(s)\, ds + B(t)v(t)\right]\right\rangle dt, \end{aligned} \tag{4.8}$$

where the first integral is understood in the Lebesgue–Stieltjes sense. Applying the unsymmetric Fubini theorem [4] to the first integral and the usual Fubini theorem to the remaining two the expression in (4.8) becomes

$$\begin{aligned} &\int_{t_0}^{t_1} \left\{\int_{t_1-2h}^{t_1-h} \lambda(t + h - t_1)^* A_1(t + h)d_t\, X(t, s) \right. \\ &\quad + \int_{t_1-2h}^{t_1-h} \dot\lambda(t + h - t_1)^* A_3(t + h)X(t, s)\, dt \\ &\quad \left. + \int_{t_1-h}^{t_1} \lambda(t - t_1)^* A_2(t)X(t, s)\, dt\right\} B(s)v(s)\, ds \\ &\quad + \int_{t_1-h}^{t_1} \lambda(t - t_1)^* B(t)v(t)\, dt. \end{aligned} \tag{4.9}$$

Thus the expression in (4.9) is the same thing as the integral term in (4.5). Extend the Borel measurable function $\dot\lambda$ as follows:

$$\dot\lambda(s) \equiv 0, \qquad s \notin [-h, 0]. \tag{4.10}$$

Hence (4.2) can be rewritten using (4.4), (4.5), (4.9), and (4.10) to obtain

$$
\begin{aligned}
&-(\bar{x}^0 - x^0) + \langle \bar{x}_{t_1} - x_{t_1}, \lambda \rangle \\
&= \int_{t_0}^{t_1} \bar{L}_u(s)^* v(s)\, ds + \int_{t_0}^{t_1} \left[\int_{t_0}^{t_1} \bar{L}_x(t)^* X(t, s)\, dt \right] B(s) v(s)\, ds \\
&\quad - \int_{t_0}^{t_1} \lambda(-h)^* X(t_1 - h, s) B(s) v(s)\, ds - \int_{t_0}^{t_1} \dot{\lambda}(s - t_1)^* B(s) v(s)\, ds \\
&\quad - \int_{t_0}^{t_1} \left\{ \int_{t_1-2h}^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_1(t + h) d_t X(t, s) \right. \\
&\quad + \int_{t_1-2h}^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_3(t + h) X(t, s)\, dt \\
&\quad \left. + \int_{t_1-h}^{t_1} \dot{\lambda}(t - t_1)^* A_2(t) X(t, s)\, dt \right\} B(s) v(s)\, ds + o(\| v \|_2) \geqslant 0, \qquad (4.11)
\end{aligned}
$$

for $v \equiv u - \bar{u}$ (v was also taken to be Borel measurable as was $\dot{\lambda}$). Define a function ψ by the following equations:

$$
\begin{aligned}
\psi(s)^* \equiv\ & \dot{\lambda}(s - t_1)^* + \lambda(-h)^* X(t_1 - h, s) \\
& - \int_{s}^{t_1} \bar{L}_x(t)^* X(t, s)\, dt + \int_{t_1-h}^{t_1} \dot{\lambda}(t - t_1)^* A_2(t) X(t, s)\, dt \\
& + \int_{t_1-2h}^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_1(t + h)\, d_t X(t, s) \\
& + \int_{t_1-2h}^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_3(t + h) X(t, s)\, dt, \qquad (4.12)
\end{aligned}
$$

for $t_0 \leqslant s \leqslant t_1$, and

$$\psi(s)^* = 0, \qquad s > t_1 \qquad (4.13)$$

Then we can write the inequality in (4.11) as

$$\int_{t_0}^{t_1} [\bar{L}_u(s)^* - \psi(s)^* B(s)] v(s)\, ds + o(\| v \|_2) \geqslant 0, \qquad (4.14)$$

for $v = u - \bar{u}$ and u an admissible controller. We conclude that

$$g(s)^* \equiv \bar{L}_u(s)^* - \psi(s)^* B(s) = 0 \qquad \text{a.e. on } [t_0, t_1]. \qquad (4.15)$$

To see that this is true just observe that the v in (4.14) is in fact an arbitrary function in $L_2([t_0, t_1], R^m)$. Thus define

$$v_\epsilon(s) = -\epsilon g(s), \qquad t_0 \leqslant s \leqslant t_1,$$

and use v_ϵ in (4.14) to obtain

$$-\epsilon \| g \|_2^2 + o(\epsilon) \geqslant 0.$$

This implies $\| g \|_2 = 0$ so that (4.15) follows at once. We now give a lemma which will complete the proof of Theorem 4.1.

LEMMA 4.1. *The function* ψ *defined by* (4.12) *and* (4.13) *satisfies the adjoint equation* (*Theorem* 4.1 (B), *condition* (ii)).

The proof of this statement involves some tedious calculations which will be deferred to an appendix.

Remark 4.2. It is noted that condition (i) of Theorem 4.1 gives

$$\bar{u}(t) = \tfrac{1}{2}[N(t)^{-1}B^*(t)\,\psi(t)] \qquad \text{a.e. on } [t_0, t_1]. \tag{4.15}$$

Since $N(t)$ is positive definite, this clearly implies

$$-L(t, \bar{x}(t), \bar{u}(t)) + \psi^*(t)B(t)\bar{u}(t) = \max_{u \in R^m}[-L(t, \bar{x}(t), u) + \psi^*(t)B(t)u] \tag{4.16}$$

a.e. on $[t_0, t_1]$. Indeed, in the geometric proof of Theorem 4.1 when we reached inequality (4.2) we could have taken a different tack and extended the proof in [12, p. 214] so that (4.16) would have resulted from our computations first and (4.15) would have been a consequence. We chose the approach we did in order to dovetail the proofs of Theorems 4.1 and 4.2. However, using the modified approach just suggested one can easily extend our methods to cover more general convex problems where

$$L(t, x, u) = f^0(t, x) + h^0(t, u)$$

and f^0, h^0 satisfy suitable growth, smoothness, and convexity assumptions [12, pp. 206]. Finally we note that the necessary conditions in Theorem 4.1 turn out to also be sufficient conditions for $(\bar{x}, \bar{u})$ to be a solution to problem (P) for retarded systems ($A_1 \equiv 0$) [9]. For neutral systems the situation is more involved. At this time we will merely observe that that if the necessary conditions of Theorem 4.1 (B) are fulfilled by the pair $(\bar{x}, \bar{u})$ and if ψ is of bounded variation, then $(\bar{x}, \bar{u})$ is optimal [2]. In many examples [2, 9], ψ does indeed turn out to be of bounded variation.

Now let us just assume that $L: R \times R^n \times R^m \to R$ is a continuous mapping with continuous first partial derivatives with respect to x and u. We further assume

$$-\infty < J(u) \equiv J(x(\cdot, \varphi, u), u) < \infty,$$

for each $u \in L_2([t_0, t_1], R^m)$, and that $J: L_2([t_0, t_1], R^m) \to R$ is continuously differentiable at each $u \in L_2([t_0, t_1], R^m)$ with

$$J'(u)(v) = \int_{t_0}^{t_1} \Big[\frac{\partial L}{\partial x}(t, x(t, \varphi, u), u(t))^* x(t, 0, v) + \frac{\partial L}{\partial u}(t, x(t, \varphi, u), u(t))^* v(t)\Big] dt, \tag{4.17}$$

$v \in L_2([t_0, t_1], R^m)$. This implicitly requires certain types of growth assumptions on L, $\partial L/\partial x$, and $\partial L/\partial u$ which we shall not make explicit [9]. We want to exhibit another derivation of necessary conditions for problem (P). These results cover the case explicitly stated in Theorem 4.1. However, the approach in Theorem 4.1 in addition to providing geometric insight, is often easier to extend to other types of problems than is the Lagrange Multiplier derivation in Theorem 4.2.

THEOREM 4.2. *Let L satisfy the conditions stated in the preceding paragraph. Let hypotheses* (H1)–(H3) *be satisfied. If* $(\bar{x}, \bar{u})$ *is a solution to problem* (P), *then the conditions* (i) *and* (ii) *of Theorem* 4.1 (B) *must be satisfied.*

Proof. Just as in the proof of Theorem 4.1 no loss of generality results in assuming $\varphi \equiv 0$ in (2.1). Since $(\bar{x}, \bar{u})$ is a solution to Problem (P), $x_{t_1}(\cdot, 0, \bar{u}) = \zeta$ (ζ is the function in (2.3)). Thus $\zeta \in \mathscr{A}_0$. Moreover, the mapping

$$u \mapsto x_{t_1}(\cdot, 0, u) - \zeta$$

maps $L_2([t_0, t_1], R^m)$ into $\mathscr{A}_0$ and

$$u \mapsto x_{t_1}(\cdot, 0, u), \qquad u \in L_2([t_0, t_1], R^m)$$

is a bounded linear operator. Thus problem (P) is that of minimizing $J(u)$ subject to

$$x_{t_1}(\cdot, 0, u) - \zeta = 0, \qquad u \in L_2([t_0, t_1], R^m).$$

In view of the results in Section 3, the Lagrange multiplier theorem [13, p. 243] applies to this constrained minimization problem. Hence there exists $\lambda \in W_2^{(1)}([-h, 0], R^n)$ such that

$$-J'(\bar{u})(v) + \langle \lambda(-h), x(t_1 - h, 0, v)\rangle + \int_{-h}^{0} \langle \dot{\lambda}(\theta), \dot{x}(t_1 + \theta, 0, v)\rangle \, d\theta \equiv 0 \tag{4.18}$$

for every $v \in L_2([t_0, t_1], R^m)$. Now $x(\cdot, 0, v)$ is just an abbreviation for the function

$$x(t, 0, v) = \begin{cases} 0, & t_0 - h \leqslant t \leqslant t_0, \\ \int_{t_0}^{t} X(t, s)B(s)v(s) \, ds, & t_0 \leqslant t \leqslant t_1. \end{cases} \tag{4.19}$$

In the proof of Theorem 4.1 inequalities (4.2) and (4.4) are combined to write

$$-J'(\bar{u})(v) + \langle \bar{x}_{t_1} - x_{t_1}, \lambda\rangle + o(\| v \|_2) \geqslant 0. \tag{4.20}$$

for each $v \in L_2([t_0, t_1], R^m)$. It is then clear that the same manipulations that led to (i) and (ii) in Theorem 4.1 yield the same result when applied to (4.18).

Appendix

In this appendix we will verify Lemma 4.1, i.e., the function ψ defined by (4.12) and (4.13) with convention (4.10) satisfies the adjoint equation (ii) in Theorem 4.1 (B). We note that (4.12) and (4.13) with convention (4.10) are equivalent to

$$\begin{aligned} \psi(s)^* = {} & \dot{\lambda}(s - t_1)^* + \lambda(-h)^* X(t_1 - h, s) - \int_s^{t_1} \bar{L}_x(t)^* X(t, s) \, dt \\ & + \int_{t_1 - h}^{t_1} \dot{\lambda}(t - t_1)^* A_2(t) X(t, s) \, dt + \dot{\lambda}(s + h - t_1)^* A_1(s + h) \\ & + \int_{s^+}^{t_1 - h} \dot{\lambda}(t + h - t_1)^* A_1(t + h) d_t X(t, s) \\ & + \int_s^{t_1 - h} \dot{\lambda}(t + h - t_1)^* A_3(t + h) X(t, s) \, dt \end{aligned} \tag{A.1}$$

for all $s \geqslant t_0$. First assuming $s \leqslant t_1 - h$ and using (2.7) and (2.8) in (A.1) we obtain

$$\begin{aligned}
\psi(s)^* = \dot\lambda(s - t_1)^* + \lambda(-h)^* \Big\{ I &+ X(t_1 - h, s + h)A_1(s + h) \\
&+ \int_s^{t_1-h} X(t_1 - h, \alpha)A_2(\alpha)\, d\alpha + \int_{s+h}^{t_1-h} X(t_1 - h, \alpha)A_3(\alpha)\, d\alpha \Big\} \\
&- \int_s^{t_1} \bar L_x(t)^* \Big\{ I + X(t, s + h)A_1(s + h) \\
&+ \int_s^t X(t, \alpha)A_2(\alpha)\, d\alpha + \int_{s+h}^t X(t, \alpha)A_3(\alpha)\, d\alpha \Big\}\, dt \\
&+ \int_{t_1-h}^{t_1} \dot\lambda(t - t_1)^* A_2(t) \Big\{ I + X(t, s + h)A_1(s + h) \\
&+ \int_s^t X(t, \alpha)A_2(\alpha)\, d\alpha + \int_{s+h}^t X(t, \alpha)A_3(\alpha)\, d\alpha \Big\}\, dt \\
&+ \dot\lambda(s + h - t_1)^* A_1(s + h) \\
&+ \int_{s^+}^{t_1-h} \dot\lambda(t + h - t_1)^* A_1(t + h) d_t \Big\{ I + X(t, s + h)A_1(s + h) \\
&+ \int_s^t X(t, \alpha)A_2(\alpha)\, d\alpha + \int_{s+h}^t X(t, \alpha)A_3(\alpha)\, d\alpha \Big\} \\
&+ \int_s^{t_1-h} \dot\lambda(t + h - t_1)^* A_3(t + h) \Big\{ I + X(t, s + h)A_1(s + h) \\
&+ \int_s^t X(t, \alpha)A_2(\alpha)\, d\alpha + \int_{s+h}^t X(t, \alpha)A_3(\alpha)\, d\alpha \Big\}\, dt. \qquad \text{(A.2)}
\end{aligned}$$

Interchanging orders of integration several times, collecting terms, and applying (2.8) we can rewrite (A.2) as

$$\begin{aligned}
\psi(s)^* = \dot\lambda(s - t_1)^* + \lambda(-h)^* &- \int_s^{t_1} \bar L_x(t)^*\, dt \\
&+ \int_{t_1-h}^{t_1} \dot\lambda(t - t_1)^* A_2(t)\, dt + \dot\lambda(s + h - t_1)^* A_1(s + h) \\
&+ \int_s^{t_1-h} \dot\lambda(t + h - t_1)^* A_3(t + h)\, dt
\end{aligned}$$

$$+ \Big\{\lambda(-h)^* X(t_1 - h, s + h) - \int_s^{t_1} \bar{L}_x(t)^* X(t, s + h)\, dt$$

$$+ \int_{t_1-h}^{t_1} \dot{\lambda}(t - t_1)^* A_2(t) X(t, s + h)\, dt$$

$$+ \int_{s^+}^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_1(t + h) d_t\, X(t, s + h)$$

$$+ \int_s^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_3(t + h) X(t, s + h)\, dt\Big\} A_1(s + h)$$

$$+ \int_s^{t_1} \Big\{\lambda(-h)^* X(t_1 - h, \alpha) - \int_s^{t_1} \bar{L}_x(t)^* X(t, \alpha)\, dt$$

$$+ \int_{t_1-h}^{t_1} \dot{\lambda}(t - t_1)^* A_2(t) X(t, \alpha)\, dt$$

$$+ \int_{s^+}^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_1(t + h) d_t\, X(t, \alpha)$$

$$+ \int_s^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_3(t + h) X(t, \alpha)\, dt\Big\} A_2(\alpha)\, d\alpha$$

$$+ \int_{s+h}^{t_1} \Big\{\lambda(-h)^* X(t_1 - h, \alpha) - \int_s^{t_1} \bar{L}_x(t)^* X(t, \alpha)\, dt$$

$$+ \int_{t_1-h}^{t_1} \dot{\lambda}(t - t_1)^* A_2(t) X(t, \alpha)\, dt$$

$$+ \int_{s^+}^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_1(t + h) d_t\, X(t, \alpha)$$

$$+ \int_s^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_3(t + h) X(t, \alpha)\, dt\Big\} A_3(\alpha)\, d\alpha. \tag{A.3}$$

Now Eq. (A.3) is the same as

$$\psi(s)^* = \dot{\lambda}(s - t_1)^* + \lambda(-h)^* - \int_s^{t_1} \bar{L}_x(t)^*\, dt$$

$$+ \Big\{\dot{\lambda}(s + h - t_1)^* + \lambda(-h)^* X(t_1 - h, s + h)$$

$$- \int_{s+h}^{t_1} \bar{L}_x(t)^* X(t, s + h)\, dt$$

$$
\begin{aligned}
&+ \int_{t_1-h}^{t_1} \lambda(t - t_1)^* A_2(t) X(t, s + h)\, dt + \dot{\lambda}(s + 2h - t_1)^* A_1(s + 2h) \\
&+ \int_{(s+h)^+}^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_1(t + h) d_t\, X(t, s + h) \\
&+ \int_{s+h}^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_3(t + h) X(t, s + h)\, dt \Big\} A_1(s + h) \\
&+ \int_{s}^{t_1} \Big\{ \lambda(-h)^* X(t_1 - h, \alpha) + \dot{\lambda}(\alpha - t_1)^* \\
&- \int_{\alpha}^{t_1} \bar{L}_x(t)^* X(t, \alpha)\, dt + \int_{t_1-h}^{t_1} \lambda(t - t_1)^* A_2(t) X(t, \alpha)\, dt \\
&+ \dot{\lambda}(\alpha + h - t_1)^* A_1(\alpha + h) \\
&+ \int_{\alpha^+}^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_1(t + h) d_t\, X(t, \alpha) \\
&+ \int_{\alpha}^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_3(t + h) X(t, \alpha)\, dt \Big\} A_2(\alpha)\, d\alpha \\
&+ \int_{s+h}^{t_1} \Big\{ \dot{\lambda}(\alpha - t_1)^* + \lambda(-h)^* X(t_1 - h, \alpha) \\
&- \int_{\alpha}^{t_1} \bar{L}_x(t)^* X(t, \alpha)\, dt \\
&+ \int_{t_1-h}^{t_1} \lambda(t - t_1)^* A_2(t) X(t, \alpha)\, dt + \dot{\lambda}(\alpha + h - t_1)^* A_1(s + h) \\
&+ \int_{\alpha^+}^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_1(t + h) d_t\, X(t, \alpha) \\
&+ \int_{\alpha}^{t_1-h} \dot{\lambda}(t + h - t_1)^* A_3(t + h) X(t, \alpha)\, dt \Big\} A_3(\alpha)\, d\alpha. \qquad \text{(A.4)}
\end{aligned}
$$

Comparing (A.4) with (A.1) we find that

$$
\begin{aligned}
\psi(s)^* = \dot{\lambda}(s - t_1)^* + \lambda(-h)^* - \int_{s}^{t_1} \bar{L}_x(t)^*\, dt + \psi(s + h)^* A_1(s + h) \\
+ \int_{s}^{t_1} \psi(\alpha)^* A_2(\alpha)\, d\alpha + \int_{s+h}^{t_1} \psi(\alpha)^* A_3(\alpha)\, d\alpha, \qquad \text{(A.5)}
\end{aligned}
$$

or, what is the same,

$$\psi(s)^* = \dot{\lambda}(s - t_1)^* + \lambda(-h)^* + \psi(s + h)^* A_1(s + h)$$
$$+ \int_{s+h}^{t_1} \psi(\alpha)^* A_3(\alpha)\, d\alpha + \int_s^{t_1} [\psi(\alpha)^* A_2(\alpha) - \bar{L}_x(\alpha)^*]\, d\alpha \quad \text{(A.6)}$$

for $s \leqslant t_1 - h$.

We turn now to the case where $t_1 - h < s \leqslant t_1$. From (A.1), (2.7), (2.8), and (4.10) we have

$$\psi(s)^* = \dot{\lambda}(s - t_1)^* - \int_s^{t_1} \bar{L}_x(t)^* \left\{I + \int_s^t X(t, \alpha) A_2(\alpha)\, d\alpha\right\} dt$$
$$+ \int_s^{t_1} \dot{\lambda}(t - t_1)^* A_2(t) \left\{I + \int_s^t X(t, \alpha) A_2(\alpha)\, d\alpha\right\} dt$$
$$= \dot{\lambda}(s - t_1)^* - \int_s^{t_1} \bar{L}_x(t)^*\, dt + \int_s^{t_1} \dot{\lambda}(t - t_1)^* A_2(t)\, dt$$
$$+ \int_s^{t_1} \left\{\int_s^{t_1} \dot{\lambda}(t - t_1)^* A_2(t) X(t, \alpha)\, dt\right.$$
$$\left. - \int_s^{t_1} \bar{L}_x(t)^* X(t, \alpha)\, dt\right\} A_2(\alpha)\, d\alpha$$
$$= \dot{\lambda}(s - t_1)^* - \int_s^{t_1} \bar{L}_x(t)^*\, dt$$
$$+ \int_s^{t_1} \left\{\dot{\lambda}(\alpha - t_1)^* - \int_\alpha^{t_1} \bar{L}_x(t)^* X(t, \alpha)\, dt\right.$$
$$\left. + \int_\alpha^{t_1} \dot{\lambda}(t - t_1)^* A_2(t) X(t, \alpha)\, dt\right\} A_2(\alpha)\, d\alpha. \quad \text{(A.7)}$$

Therefore, we have for $t_1 - h < s \leqslant t_1$

$$\psi(s)^* = \dot{\lambda}(s - t_1)^* + \int_s^{t_1} [\psi(\alpha)^* A_2(\alpha) - \bar{L}_x(\alpha)^*]\, d\alpha. \quad \text{(A.8)}$$

Now define

$$\mu(s)^* \equiv \begin{cases} \dot{\lambda}(s - t_1)^* + \lambda(-h)^*, & s \leqslant t_1 - h, \\ \dot{\lambda}(s - t_1)^*, & s > t_1 - h, \end{cases} \quad \text{(A.9)}$$

and

$$\psi(s)^* = 0 \qquad \text{if} \quad s > t_1 . \tag{A.10}$$

Applying (A.9) and (A.10) in (A.6) and (A.8) we have that

$$\psi(s)^* = \mu(s)^* + \psi(s+h)^* A_1(s+h) + \int_{s+h}^{t_1} \psi(\alpha)^* A_3(\alpha)\, d\alpha$$
$$+ \int_s^{t_1} [\psi(\alpha)^* A_2(\alpha) - \bar{L}_x(\alpha)^*]\, d\alpha$$

for $s \in [t_0, t_1]$. This completes the proof of Lemma 4.1.

References

1. H. T. Banks, Control of functional differential equations with function space boundary conditions, *in* "Proceedings of the Park City Differential Equations Symposium," Academic Press, New York, 1972.
2. H. T. Banks and G. A. Kent, Control of functional differential equations of retarded and neutral type to target sets in function space, *SIAM J. Control* **10** (1972), 567–594.
3. R. Bellman and K. L. Cooke, "Differential Difference Equations," Academic Press, New York, 1963.
4. R. H. Cameron and W. T. Martin, An unsymmetric Fubini theorem, *Bull. Amer. Math. Soc.* **47** (1941), 121–125.
5. N. Dunford and J. T. Schwartz, "Linear Operators," Part I, Interscience, New York, 1958.
6. A. Halanay, On the controllability of linear difference-differential systems, *in* "Math. Systems Theory and Economics" II (H. W. Kuhn and G. P. Szegö, Eds.), Lecture Notes in Operations Research and Mathematical Economics No. 12, Springer-Verlag, New York, 1969.
7. J. K. Hale and K. R. Meyer, A class of functional equations of neutral type, *Mem. Amer. Math. Soc.* **76** (1967), 1–65.
8. M. Q. Jacobs, An optimization problem for an n-th order scalar neutral functional differential equation with functional side conditions, *in* "Proceedings of the Park City Differential Equations Symposium," Academic Press, New York, 1972.
9. M. Q. Jacobs and T. J. Kao, An optimum settling problem for time lag systems, *J. Math. Anal. Appl.* **40** (1972), 1–21.
10. F. M. Kirillova and S. V. Curakova, The controllability problem for linear systems with aftereffect, *Differencial'nye Uravnenija* **3** (1967), 436–445.
11. V. L. Klee, The support property of a convex set in a linear normed space, *Duke Math. J.* **15** (1948), 767–772.
12. E. B. Lee and L. Markus, "Foundations of Optimal Control Theory," Wiley, New York, 1967.
13. D. G. Luenberger, "Optimization by Vector Space Methods," Wiley, New York, 1969.
14. E. H. Moore, General analysis. I, *Memoir Amer. Phil. Soc.* **1** (1935), 170–220.

15. L. W. Neustadt, A general theory of extremals, *J. Comput. System Sci.* **3** (1969), 57–92.
16. R. Penrose, A generalized inverse for matrices, *Proc. Cambridge Phil. Soc.* **51** (1955), 406–413.
17. A. E. Taylor, "General Theory of Functions and Integration," Blaisdell, New York, 1965.
18. L. Weiss, On the controllability of delay-differential systems, *SIAM J. Control* **5** (1967), 575–587.
19. L. Weiss, Lectures on controllability and observability, Technical Note BN-590, University of Maryland, 1969.
20. R. B. Zmood, On Euclidean space and function space controllability of control systems with delay, Technical Report, University of Michigan, June, 1971.